

INFORMATION

AGENDA ITEM: Statewide Information Databases – California Digital Library's Government Information Initiative

GENERAL OVERALL PROGRAM UPDATES:

CURRENT STATUS: California state government agencies continue to provide more and more of their information in digital format via the World Wide Web. Long-term retention of this state government information, however, is not guaranteed. There is no assurance that information available on a state agency Web site today will be available tomorrow. In addition, information on different state agency Web sites is distributed in many different formats with a variety of access methods, often making it difficult for Californians to find and use the information they need.

In order to remedy this situation, the California Digital Library of the University of California (CDL) has created a concept known as the Government Information Initiative, a proposal to develop a feasible way of acquiring and providing continuous and permanent access to government information available through California state agency Web sites, and to make this access available continuously to all Californians.

The California State Library (CSL) and CDL have signed an interagency agreement whereby CSL agrees to fund and CDL agrees to facilitate the California Data Warehouse, the first phase of the Government Information Initiative. Library of California state operations funds are being used for this agreement, and this will be one of the first projects under the collaborative efforts between the Library of California and the California Digital Library.

CDL will develop a single, easy-to-use, World Wide Web interface through which users can easily locate, combine, use, and manipulate data from a variety of sources, without the need to identify the original data source. This system is the California Data Warehouse: a collection of disparate data from multiple sources, united into a single collection and format that can be easily searched and used by a variety of users. The Warehouse will be maintained by CDL.

Initially the contents of the California Data Warehouse will focus on the social sciences. These data will come from a variety of sources (primarily California state agencies, but will also include local and federal data) that are already publicly available but in disparate formats. The Warehouse will include data on population, housing, health, crime, economic, education, environment and other topics. Most importantly, the social science data that is being targeted are used by a wide variety of people across the state, whether they are patrons in a public library, clients of a corporate library, high school students, or researchers from the

University of California. Given this wide range of potential users, the California Data Warehouse will serve a very broad audience throughout the state.

The \$318,374 in Library of California state operations funds will be matched by a comparable amount in CDL funding to make the resources, created by taxpayer dollars, available to Californians as long as they are valuable state resources to share.

Activities included in the agreement comprise the first phase of the implementation of the project. The first phase includes designing the technical structure, architecture, design, and user interface of the California Data Warehouse. CDL staff also will work with State Library staff to identify which agencies are producing data and what format the data are in, and to obtain and load sample data from the agencies into the Warehouse. The second phase of the project will focus on adding contents to the database and developing user tools. The State Librarian is considering awarding a Library Services and Technology Act (LSTA) grant to supplement CDL's funding of the second phase.

The California Research Bureau (CRB) at the California State Library is very excited about this project and about it becoming both a CDL and Library of California joint effort. In fact, over the past couple of years CRB researchers and librarians have discussed the need for such an undertaking, in part because they were hearing about that need from policy makers within state government, particularly the Health and Human Services Agency, and because they believe that CDL has the technical and project management capabilities to undertake such a project and make it a success.

In reviewing this proposed project, John Jewell, the CRB's Assistant Director for Information Services, stated:

"It would seem to me that in this early stage having both CDL and LoC perceived as helping move such an enterprise forward would be advantageous to both programs. CDL obviously would receive much credit and praise for the implementation and ongoing operations. LoC would be able to showcase a major unique resource for the general public and the library community, later provide training for non-UC library staff, and demonstrate the unique partnering that these two programs could generate."

RELATED ISSUE TO COME BEFORE THE BOARD IN THE FUTURE: Update on the California Data Warehouse.

Relevant Committee: Support Services
Staff Liaison: Tom Andersen